Detection of multicomponent signals: effect of difference in level between components

Dubois Françoise, Sabine Meunier, Rabau Guy, Poisson Franck, Guyader Gaël

Detection of multicomponent signals: Effect of difference in level between components

Françoise Dubois, Sabine Meunier, Guy Rabau

Laboratoire de Mécanique et d'Acoustique, Unité Propre de Recherche 7051,

Centre National de la Recherche Scientifique, 31 chemin Joseph-Aiguier, 13402 Marseille, France

dubois@lma.cnrs-mrs.fr, meunier@lma.cnrs-mrs.fr, dubois@lma.cnrs-mrs.fr

Franck Poisson

Société Nationale des Chemins de fer Français, 45 rue de Londres, 75379 Paris, France

franck.poisson@sncf.fr

Gaël Guyader

Technocentre Renault,, 1 avenue du Golf, 78288 Guyancourt, France

gael.guyader@renault.com

Abstract: The detection of multicomponent signals for which the components are not equidetectable is precisely investigated as a function of the level difference $\Delta L_{i/j}$ between components. The detection thresholds are determined for a seven-tone complex signal with random starting phases masked by white noise. Level differences between the components are examined. A model for non-equidetectable conditions based on the statistical summation model is described. The improvement in detection is calculated from the level difference between components that is related to the thresholds for single components. The model predictions are in accordance with the experimental results.

PACS numbers: 43.66.Dc

Running title: Detection of multicomponent signals

1. Introduction

When trying to detect a signal in background noise, the listener is assumed to use a filter whose center frequency is equal to that of the signal. Threshold is assumed to correspond to a certain signal-to-noise ratio at the output of the filter. This set of assumptions has come to be known as the *power spectrum model* of masking (for a review see Green, 1958). Gässler (1954) and Marill (1956) showed that the power spectrum model was also applicable to multicomponent signals: they found that the detection of a complex tone is based on the detection of its most easily detectable component. According to these authors, the detection of multicomponent signals is only based on the signal energy falling into one frequency band. However, Green (1958) found that a pair of signals widely separated in frequency was better detected than its more detectable member. The improvement in detection for multicomponent signals is defined as the difference in level between the threshold of each component when presented in isolation and the masked threshold of the complex expressed as level per component. Later experiments actually confirmed the improvement in detection for multicomponent signals (Spiegel, 1979; Langhans and Kohlrausch, 1992). Multiple widely spaced sinusoidal components presented in a background noise are better detected than any of the individual components.

Two models have been proposed to predict the masked threshold of a complex signal, both based on the detection of each individual component when presented alone in the same masker: the decision-threshold model and the statistical summation model. The statistical summation model (Green and Swets, 1974) postulates that the energy within each frequency-selective channel is transformed into a Gaussian distributed decision variable. Thus, a detectability index d_i' is associated to each channel *i*. The overall detectability index d_n' is the integration of the individual

d_i'. Thus, according to the statistical summation model, multiple channels are integrated to form a single basis for detection. An increase in the number of stimulated channels raises the overall value of the detectability index, and consequently the detection is improved. Contrary to the statistical summation model, the decision-threshold model predicts that the overall decision is a function of the individual decision in each channel. The overall performance in detection is based on multiple observations from the different stimulated channels. Each additional observation (channel) leads to another independent detection opportunity, and consequently the detection is improved. Both models are valid provided the observations in the different channels are independent.

The two models are often indistinguishable because both predict an improvement in detection. Buus et al. (1986) measured psychometric functions for three pure tones and an 18-tone complex masked by a uniformly masking noise. In order to distinguish between the two models, Buus et al. (1986) presented either each stimulus in different sessions (fixed condition) or all the stimuli randomly within one session (random condition). According to the decision-threshold model, the probability that in each stimulated channel the decision variable exceeds a fixed percentage of correct responses related to the threshold is the same in both conditions. Consequently, the multicomponent signal threshold should be unchanged. Contrary to the decision-threshold model, the statistical summation model predicts an increase in the variance of the decision variable in the random condition, and consequently, an increase in thresholds. Buus et al. (1986) measured thresholds that were higher in the random condition than in the fixed condition, indicating that the statistical summation model is in good agreement with their experimental data. According to the statistical summation model, the overall detectability index d_n' is equal to

$d'_n = \sqrt{\sum_{i=1}^{n} d'^2_i}$. In the case of equidetectable components, the individual detectability indexes d_i' are equal and $d'_n = \sqrt{n}d'_i$. As d' is proportional to the signal intensity, the statistical summation model predicts that the detection is improved by 5 log(n) dB : $T_i - T_{i/n} = 5 * \log(n)$ (see next section for demonstration). $T_{i/n}$ is the level of the i^{th} component in the multicomponent signal at its masked threshold and T_i is the threshold of the i^{th} component when presented alone in the masker.

Few authors have studied the influence of the difference in level between components on masked thresholds. Bacon et al. (2002) measured thresholds of triplets presented at equal physical levels, masked by unmodulated or modulated narrow bands of noise, centered at the signal frequencies. They compared their results with those of Grose and Hall (1997) who measured thresholds for multicomponent signals presented at equal physical levels. Bacon et al. (2002) found that the improvement in thresholds was not influenced by the assumption of equidetectability. However, this conclusion was tempered by the fact that the difference between the single-component thresholds was probably less than a few dB for the frequencies tested. So, Bacon et al. (2002) emphasized the importance of level differences between components to study the detection of complex tones. Van den Brink and Houtgast (1990) measured the thresholds for three types of signals masked by white or pink noise. Each component in each 1/3-oct band covered by the complex signal was fixed at the same signal-to-noise ratio. So, as they admitted, the signals were not exactly - but only approximately - equidetectable. A 4 log(n) rule (not a 5log(n) rule) seemed to predict the experimental results.

In the present study, the detection of multicomponent signals is accurately investigated as a function of the level difference between components. The issue of interest is the extension of the statistical summation model to non-equidetectable components.

2. The statistical summation model for non-equidetectable components

According to the statistical summation model, if all the components of the complex are in different critical bands and equidetectable, the individual values of d_i' are equal and the combined value of *d'*, denoted d'_n, increases as the square root of *n.* In the case of non-equidetectable components, a difference in level $\Delta L_{i/j}$ is introduced in order to examine the relative contribution to global detection of each complex signal component. $\Delta L_{i/j}$ is defined as the difference between L_i and L_j, which are the levels of the i^{th} and the j^{th} components of the n-component signal, expressed relative to their individual threshold T_i and T_j. $L_{i/n}$ and $L_{j/n}$ are the levels of the ith and the jth components of the n-component signal.

$$\Delta L_{i/j} = L_i - L_j = \left(L_{i/n} - T_i\right) - \left(L_{j/n} - T_j\right) \qquad \textbf{(E1)}$$

It is known that the psychometric functions for the detection of pure tones in noise and in quiet are well described by $d'=kI^{\alpha}$, where I is the signal power and $\alpha \approx 1$ (Egan et al., 1969). Consequently, the difference between the detectability index $d'_{i/n}$ for the i^{th} component in the n-component signal and the detectability index $d'_{i(thr)}$ at threshold when the same component is presented individually can be related to the difference in level:

$$10*\log(d'_{i/n}) - 10*\log(d'_{i(thr)}) = L_{i/n} - T_i \qquad \textbf{(E2a)}$$

$$\left[10*\log(d'_{i/n}) - 10*\log(d'_{i(thr)})\right] - \left[10*\log(d'_{j/n}) - 10*\log(d'_{j(thr)})\right] = \Delta L_{i/j} \qquad \textbf{(E2b)}$$

Given that at threshold $d'_{i(thr)}$ and $d'_{j(thr)}$ are equal, the formula reduces to:

$$d'_{i/n} = 10^{\frac{L_{i/j}}{10}} d'_{j/n} \qquad \textbf{(E3)}$$

The overall detectability index, that is, the combination of the individual values of $d_{i/n}$', expressed via the single detectability index $d'_{j/n}$, becomes:

$$d'_n = \sqrt{\sum_{i=1}^{n} \left(10^{\frac{L_{i/j}}{10}}\right)^2} \; d'_{j/n} \qquad \textbf{(E4)}$$

Thus, if an optimal weighting of the channel is used, $d'_n = A \cdot d'_{j/n}$ with

$$A = \sqrt{\sum_{i=1}^{n} \left(10^{\frac{L_{i/j}}{10}}\right)^2} \qquad \textbf{(E5)}$$

$$10\log d'_{j/n} = 10\log k + L_j \text{ and } 10\log d'_n = 10\log\left(A \cdot d'_{j/n}\right) = 10\log A + 10\log k + L_{j/n} \qquad \textbf{(E6)}$$

Given that at threshold d'_n is equal to $d_{j/n}$': $T_j = 10\log A + T_{j/n}$ **(E7)**

Consequently, the improvement in detection for a non-equidetectable multicomponent signal is related to the level difference $\Delta L_{i/j}$ between components by the relation:

$$T_j - T_{j/n} = 10 * \log \sqrt{\sum_{i=1}^{n} \left(10^{\frac{L_{i/j}}{10}}\right)^2} \qquad \textbf{(E8)}$$

The purpose of the experiment presented here was to measure masked thresholds for non-equidetectable complex signals to determine precisely the extent of improvement in detection, as a function of the level difference $\Delta L_{i/j}$ between components, and then to test the statistical summation model for non-equidetectable multicomponents with an optimal weighting of the channels.

3. Detection thresholds for complex signals with differences in level between components

3.1. Method

Detection thresholds were determined for a seven-tone (80, 160, 320, 640, 1280, 2560, 5120 Hz) complex signal, with random starting phases masked by white noise. The variation of the level differences between the components was examined via the parameter $\Delta L_{i/j}$, as defined in paragraph 2. The reference component j is the fourth component of the complex signal, at 640 Hz. In condition 1, six out of the seven components (1, 2, 3, 5, 6 and 7) were decreased by $\Delta L_{i/4}$, which was equal to 0 (equidetectable condition), 5, 10, 15, 20 and 25 dB. In condition 2, four out of the seven components (1, 2, 6 and 7) were decreased by $\Delta L_{i/4}$, which was equal to 0 (equidetectable condition), 5, 10 and 15 dB. The masker was presented at a level of 55 dB SL. The headphone was calibrated using a dummy head (Head Acoustics HRS II-2) so that levels were defined as sound pressure level (SPL) near the eardrum.

Masker and signals were generated in an RPvds (TDT) circuit including 20 ms cosine-squared rise/fall ramps. They were all 500 ms in duration. The stimuli were played out of a real-time processor (RP2, TDT), then passed through programmable attenuators (TDT PA5) and a headphone driver (TDT HB7) before being presented to the listener over a Sennheiser HD650 headset. The stimuli were presented diotically to the listeners. The listeners were seated in a double-walled sound attenuating booth and gave their responses via a hand-held response box.

Thresholds were measured using an adaptive three-interval forced-choice procedure (3 IFC) incorporating a three-down one-up stepping rule that estimated the 79.4% correct point on the psychometric function. The listening intervals were indicated by a number (1, 2 or 3) on the listener's monitor, and feedback was provided after each response. Level was initially adjusted in steps of 5 dB, and reduced to 2 dB after the second reversal. The track continued for a total of twelve reversals, and the associated threshold estimate was computed as the average signal level at the last eight reversals. Thresholds were collected blocked by condition. In each block, $\Delta L_{i/4}$ was fixed. The order of block presentation was randomized and differed from one listener to the other. Two threshold estimates were collected for each $\Delta L_{i/4}$ and averaged to generate a final threshold estimate. If the standard deviation of that average was greater than 3 dB, an additional estimate was obtained and included in the average.

Nine subjects (2 females, 7 males) ranging in age from 22 to 51 years (mean of 29 years) participated in the experiment. All had absolute thresholds of 15 dB HL or less at octave frequencies between 125 and 8000 Hz.

3.2. Results

The improvement in detection was calculated for each subject. The results were consistent across subjects. The group's mean results are shown in figure 1.a. The improvement in detection is plotted for condition 1 (circles) and condition 2 (squares) as a function of $\Delta L_{i/4}$ dB. The predicted masked thresholds calculated by applying equation (E8) are presented in figure 1.a as solid line for condition 1 and dashed line for condition 2. When the components of the complex signal are equally detectable ($\Delta L_{i/4}$=0), the improvement in detection is equal to 4.22 dB, as predicted by the 5log(n) rule. In figure 1.a, for condition 1, it can be seen that as the level difference $\Delta L_{i/4}$ increases above 10 dB the improvement in detection tends towards zero. When $\Delta L_{i/4}$ is equal to 5 dB, an improvement in detection midway between 5log(n) and zero is observed. For condition 2, it can be seen that when the components of the signal are not equally detectable, the improvement in detection tends towards a constant value. This value is close to the 5log(3) dB threshold improvement, as predicted by the statistical summation model for a 3-component equally-detectable signal. A repeated-measures analysis of variance (ANOVA) revealed a significant effect of the level difference $\Delta L_{i/j}$ in both cases (F(N=9, dl=5)=20.63, p<0.0001 and F(N=9, dl=3)=5.08, p<0.01). For condition 1, the pearson correlation between measured and predicted was highly significant (R2=0.94, p=0.0016). For condition 2, the correlation was not significant (R2=0.77, p=0.12); however, only four points were used to calculate the pearson correlation. If the predicted values were plotted against the measured values for both conditions (fig. 1.b), a linear relationship was observed with a highly significant correlation (R2=0.94, p<0.001).

3.3. Discussion

The results for $\Delta L_{i/4}$=0 are similar to those reported previously in literature for equally detectable signals. Each component of the seven-tone complex signal is treated as an independent source of information. For $\Delta L_{i/4}$=0, at signal threshold, the individual values of d_i' for the lower components tend towards zero as $\Delta L_{i/4}$ increases. Therefore, the overall sensitivity of the multitonal complex (d_n') tends toward the index detectability of the higher components, and thus only these components play a role in the detection. The improvement in detection tends to the improvement in detection found for a signal containing only the prominent components.

The most interesting aspect of the present results is the improvement in detection observed in the midway condition. Detection is improved by a factor of less than 5log(7) or 5log(3), but still it is improved. The 5log(n) rule is not valid. Indeed, the d_i' of the softer components contribute less to the overall d_n' than in the equally detectable condition. The data confirm previous results, obtained by Buus and Grose (2009). They measured thresholds for 5-component complex signals masked by 15-Hz wide bands of Gaussian noise. Each tone was presented at equal level in dB SL (normalized) or at equal level in dB SPL (equal). They showed that the 5log(n) rule seems to fail for equally intense components whereas it is well established for equally detectable components: they measured a 3.3 dB threshold improvement (5log(5)=3.49) for the five-component signal in the normalized condition, when each tone was individually adjusted in amplitude based on the detection threshold, and only 1.5 dB in the equal condition.

The extension of the statistical summation model to non-equitectable components allows to predict quite well measured values. However, the correlation between the predicted and the measured values was not good in condition 2. The condition $\Delta L_{i/4}$ = -5 dB is the one for which observed and predicted values are the most different. But a t-test did not indicate significant difference between the observed and the predicted thresholds for $\Delta L_{i/4}$ = -5 dB in both conditions (t_8= 1.26, p=0.025 and t_8=1.51, p=0.025, for conditions 1 and 2 respectively). It will however be important to study more specifically the thresholds for values of $\Delta L_{i/4}$ around -5dB in order to refine the results.

4. Summary and conclusions

A model for non-equidetectable multicomponent signals was described. An experiment was conducted wherein the improvement in detection was measured for a seven-tone signal with differences in level between components. The results can be summarized as follows: (1) Firstly, for the equidetectable signal, the result is consistent with the literature results and also with the statistical summation model prediction results. (2) When the level difference $\Delta L_{i/j}$ between components is large, only the most prominent components play a role in the detection. (3) A model was proposed that makes it possible to predict the improvement in detection in the case of non-equidetectable components. These data have theoretical as well as practical implications. The results of this experiment corroborate established results concerning the integration of signal energy from across the spectrum, and extend them to the specific configuration wherein components of the multi-tone signal exhibit level differences. From a practical point of view, the data may provide a basis to predict masking effects in conditions closer to ecologically significant conditions that are typically more complex.

References and links

Bacon, S. P., Grimault N., and Lee J. (2002). "Spectral integration in bands of modulated or unmodulated noise", J. Acoust. Soc. Amer. 112 (2), 219-226.

Buus, S., Schorer E., Florentine M., and Zwicker E. (1986). "Decision rules in detection of simple and complex tones", J. Acoust. Soc. Amer. 80 (6), 1646-1657.

Buus, E., and Grose, J.H. (2009). "Spectral integration under conditions of comodulated masking release", J. Acoust. Soc. Amer., 125 (3), 1612-1621.

Egan J. P., Lindner W. A. and McFadden D. (1969). "Masking-level differences and the form of the psychometric function", Perception and Psychophysics 6 (4), 209-215.

Gässler, G. (1954). "über die Hörschwelle für Schallereignisse mit verschieden breitem Frequenzspectrum", Acustica 4, Beih. 1, 408-414.

Green, D. M. (1958). "Detection of multicomponent signals in noise", J. Acoust. Soc. Amer., 30 (10), 904-911.

Green, D.M., and Swets, J.A., (1974). "Signal Detection Theory and Psychophysics", New York: Wiley.

Grose, J. H. and Hall J. W. III, (1997). "Multiband detection of energy fluctuations", J. Acoust. Soc. Amer. 102 (2), 1088-1096.

Langhans, A. and Kohlrausch, A. (1992). "Spectral integration of broadband signals in diotic and dichotic masking experiments", J. Audio Eng. Soc. 91 (1), 317-326.

Marill, T. (1956). "Detection theory and psychophysics", Tech. Rep. No. 319, Research Laboratory of Electronics, MIT, Cambridge, MA.

Spiegel, M. F. (1979). "The range of spectral integration", J. Acoust. Soc. Amer. 66 (5), 1356-1363.

Van den brink, W.A.C., and Houtgast, T. (1990). "Efficient across-frequency integration in short-signal detection," J. Acoust. Soc. Amer. 87 (1), 284-291.

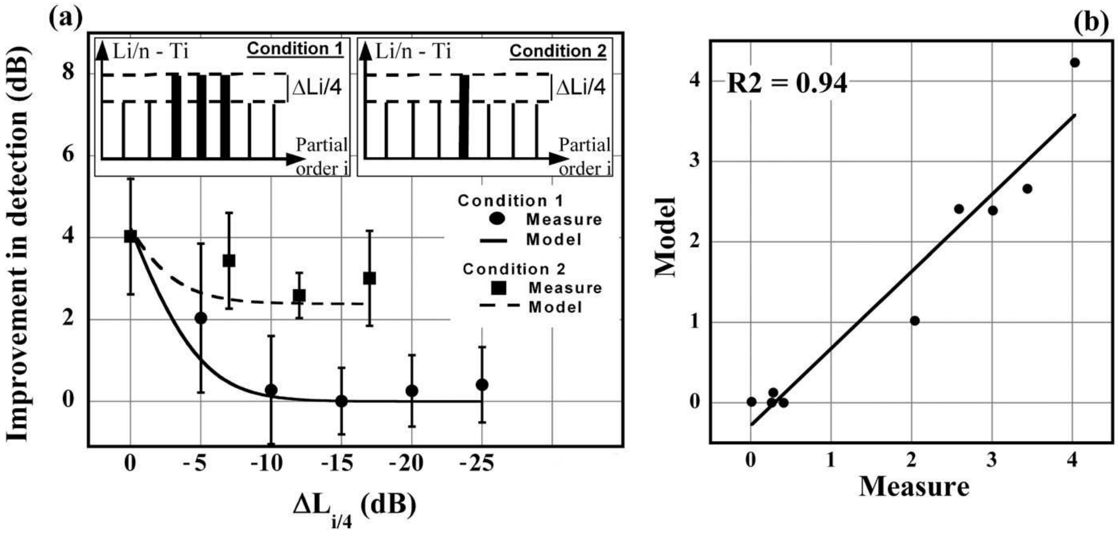

Figure 1: a) The extent of improvement in detection averaged across subjects as a function of the signal conditions is represented. In condition 1 (circles) the signal was equally detectable (0 dB) and then, the 640Hz component was increased by 5, 10, 15, 20, 25 dB. The solid line illustrates the predicted values from the statistical summation model for condition 1. In condition 2 (squares), the signal was equally detectable (0 dB) and then, the 320, 640 and 1280Hz components were increased by 5, 10, 15 dB. The squares were slightly moved to the right to make it easier to observe. The standard deviations are represented by the error bars. b) The predicted improvements in detection from the statistical summation model as a function of the measured improvements in detection is represented for conditions 1 and 2. R2 is shown on the upper right of the graph.